

3D Landmarking in Multiexpression Face Analysis:
3 A Preliminary Study on Eyebrows and Mouth

Original

3D Landmarking in Multiexpression Face Analysis: 3 A Preliminary Study on Eyebrows and Mouth / Vezzetti, Enrico; Marcolin, Federica. - In: AESTHETIC PLASTIC SURGERY. - ISSN 0364-216X. - (2014). [10.1007/s00266-014-0334-2]

Availability:

This version is available at: 11583/2544542 since: 2017-03-28T15:51:49Z

Publisher:

Springer

Published

DOI:10.1007/s00266-014-0334-2

Terms of use:

openAccess

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

3D landmarking in multi-expression face analysis: a preliminary study on eyebrows and mouth

Enrico Vezzetti, Federica Marcolin¹, Giulia Fracastoro

Department of Management and Production Engineering
Polytechnic University of Turin (Politecnico di Torino), Italy

Abstract

The application of 3D facial analysis and landmarking algorithms in the field of maxillofacial surgery and other medical applications, such as diagnosis of diseases by facial anomalies and dysmorphies, has gained a lot of attention nowadays. In a previous work, we used a geometric approach to automatically extract some three-dimensional face key-points called landmarks, working in the Differential Geometry domain, through the coefficients of the Fundamental Forms, the Principal Curvatures, Mean and Gaussian Curvatures, derivatives, Shape and Curvedness Indexes, and Tangent map. In this paper we describe the extension of our previous landmarking algorithm, which is now able to extract eyebrows and mouth landmarks, using both old and new meshes. The algorithm has been tested both on our face database and on the public Bosphorus one.

The choice of working separately with mouth and eyebrows as a detached study relies to the role that these parts have in facial expressions. In fact, since the mouth is the part of the face which moves the most and mainly affects facial expressions, extracting mouth landmarks from various face poses means that the new developed algorithm is pose-independent.

Keywords

Soft-Tissue Landmarks; Differential Geometry; 3D scanner; 3D face; Face Morphology; Medical Diagnosis.

1. Introduction

Recently, many automatic landmark extraction algorithms have been implemented on 3D medical images. In their various publications, Alker *et al.*, Frantz *et al.*, Wörz *et al.*, and Rohr *et al.* proposed some multi-step differential procedures for subvoxel localization of three dimensional landmarks, addressing the problem of choosing an optimal neighbourhood for a region-of-interest (ROI) around face landmarks (Frantz *et al.*, 1998, 1999; Perakis *et al.*, 2009). They introduced an approach for the localization of 3D anatomical point landmarks based on deformable models. To model the surface at a landmark, they used quadric surfaces combined with global deformations (Alker *et al.*, 2001; Frantz *et al.*, 2000). Then they proposed to fit 3D parametric intensity models to 3D images, using an analytic intensity model based on the Gaussian error function combined with 3D rigid transformations and deformations to efficiently model anatomical parts (Wörz and Rohr, 2005). Finally they introduced a brand new multi-step approach to improve 3D anatomical landmarks localization in tomographic images (Frantz *et al.*, 2005).

¹ Corresponding author: federica.marcolin@polito.it

Romero *et al.* compared several approaches that use graph matching and cascade filtering for landmark localization in 3D face data; for the first method, they applied the structural graph matching algorithm relaxation-by-elimination using a simple distance-to-local-plane node property and a Euclidean-distance arc property. A second method uses pose-invariant feature descriptors embedded into a cascade filter to localize the nose tip. After that, local graph matching is applied to localize the inner eye corners (Romero and Pears, 2009a). Then they described and evaluated two pose-invariant point-pair descriptors, i.e. spin image and one derived from an implicit radial basis function, which encoded 3D shape between a pair of 3D points. They showed how these descriptors were able to identify the nose-tip and the eye-corner of a human face simultaneously in six promising landmark localization systems (Romero and Pears, 2009b).

Ruiz *et al.* (2008) presented an algorithm for automatic localization of landmarks on 3D faces; an Active Shape Model (ASM) is used as a statistical joint location model for configurations of facial features. The ASM is adapted to individual faces through a guided search whereby landmark specific shape index models are matched to local surface patches. Similarly, Sang-Jun and Dong-Won (2008) applied the Active Shape Models to extract the position of the eyes, the nose and the mouth. Salah and Akarun (2006b) proposed a method for facial landmarks localization that relies on unsupervised modelling of landmark features obtained through different Gabor filter channels. They also showed that using directly three-dimensional features is more promising than illumination correction with the help of 3D (Salah and Akarun, 2006a). Then, the same work group inspected shortcomings of existing known-in-literature landmark detection approaches and compared several methods for performing automatic landmark localization on ‘almost frontal positioned’ faces in different scales. They employed two new methods for analyzing facial features in coarse and fine scales successively. A mixture of factor analyzers is used in the first method to learn Gabor filter outputs on a coarse scale. The second method is a template matching of Discrete Cosine Transform (DCT) features (Salah *et al.*, 2007). Later, Dibeklioglu *et al.* (2008) assessed a fully automatic 3D facial landmark localization algorithm based on an accurate statistical modelling of the facial traits. Finally, several 3D face recognizers were evaluated on the Bosphorus database, gathered for studies on facial expression, by the same research team. The authors provided identification results of three 3D face recognition algorithms, i.e., generic face template-based ICP approach, one-to-all ICP approach, and depth image-based Principal Component Analysis (PCA). In addition, they incorporated bi-dimensional texture classifiers in a fusion setting (Alyüz *et al.*, 2008). D’Hose *et al.* (2007) presented a landmarks localization method on 3D faces using Gabor wavelets to extract the curvature of the 3D faces, which is then used for performing a coarse detection of landmarks.

More recently, other authors worked on 3D landmark extraction. Among them, we cite Creusot *et al.* (2013), whose study dealt with generic key point detection process for 3D meshes, based on machine learning. It is then employed in a three-dimensional facial landmarking system in order to improve upon existing heuristic approaches. Zhang *et al.* (2013) proposed an automatic landmark localization approach for a 3D face, which does not need any face image and texture information. The approach employs a new local descriptor named *HoSNI* (Histogram of Shape Normal Information) to characterize local shape around a point on the three-dimensional face.

Our previous studies by Vezzetti and Marcolin (2012) and Vezzetti *et al.* (2013) relied on analyses of face morphology only through geometrical features, with the aim of formalizing the geometry of the most famous anatomical face landmarks, using morphological operators such as derivatives, shape and curvedness indexes, mean curvature, Gaussian curvature, the coefficients of the first and second fundamental forms e , f , g , E , F , and G , and the tangent map T . These works show how face analysis is possible only using geometrical descriptions of facial features, obtaining an accurate landmarks localization or, at least, a precise identification of the zones in which surely the

landmarks lie in. While in these previous work landmarks in exam were nose and eye ones, in this work the aim is to focus the attention on other facial points with a core feature that distinguishes them from the previous. This feature consists of the substantial movement the landmarks in question are subject to when faces change expression. Unlike nose and eye edges, mouth and eyebrows are consistently at the mercy of facial movements due to pose changes.

This work has the objective of providing a preliminary formalization of mouth and eyebrows 3D landmarking, thus designing an extraction algorithm that could be considered pose-independent.

2. The proposed method

A facial landmark is a point which all faces join and has a particular biological meaning. In human face up to fifty-nine soft-tissue landmarks, lying on the skin, could be collected, but the most famous and known ones are nearly twenty, as shown in Figure 1.

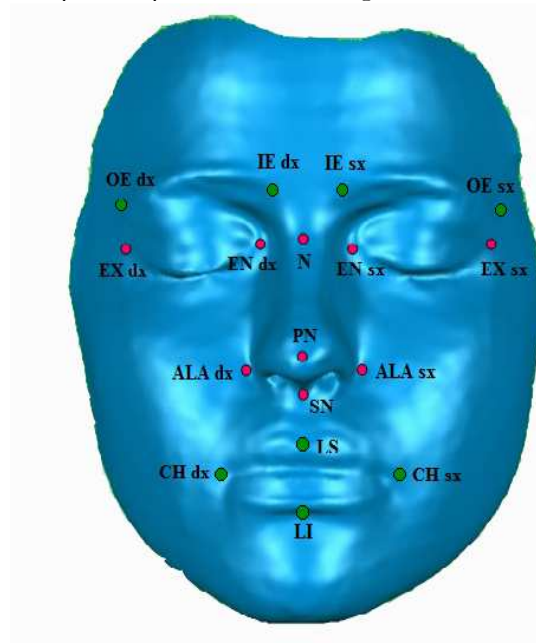


Figure 1. The landmarks detected by our proposed method. The pink are the ones previously detected: PN-pronasal, SN-subnasal, AL-alae, N-nasion, EN-endocanthions, EX-exocanthions. The green are the ones detected with the new algorithm: CH-cheilions, LS-labrum superior, LI-labrum inferior, OE-outer eyebrows, IE-inner eyebrows (Calignano and Vezzetti, 2010).

In our previous works (Vezzetti and Marcolin, 2012) and (Vezzetti *et al.*, 2013), some geometric descriptors have been used in order to detect and extract the landmarks. For length reasons, we do not report here the detailed definitions of the employed descriptors, but we refer to the ones provided in our previous works, also concerning the parameterization of the surface.

A landmark is here identified as a point of the grid with a particular biological and geometrical meaning, uniquely described by its coordinates as explained above. A geometrical formalization of each landmark belonging to the mouth and to the eyebrows, one at a time, is presented below. The localization and extraction processes used in the algorithm to detect the landmarks are explained and graphically represented.

2.1 IE- inner eyebrow

The *inner eyebrows* are the points which lie at the connection between the nose bone and the eyebrows themselves. As happened for the eye landmarks localization, even the extraction of the *inner eyebrows* is done using the coefficient of the second fundamental form g , which well describes the eye region, starting from the coordinates of the *endocanthions* previously found. The process here developed is very similar to the one used for the detection of the other points of the eyes: firstly, the *inner eyebrows* candidate points are the ones which have parametric coordinates, respectively for the right and the left eyebrow, in the ranges:

$$\begin{aligned} u_{IEr} &\in [u_{ENr} + 3, u_{ENr} + 7], & v_{IEr} &\in [v_{ENr} + 5, v_{ENr} + 20], \\ u_{IEl} &\in [u_{ENl} - 7, u_{ENl} - 3], & v_{IEl} &\in [v_{ENl} + 5, v_{ENl} + 20], \end{aligned}$$

where u_{EN} and v_{EN} are the u and v coordinates of the left or right *endocanthions*. The choice of these coefficients is not random, but it has been done by studying approximately 45 real human faces with the help of a nine-years experienced maxillofacial surgeon working at the hospital San Giovanni Battista "Molinette" of Turin, in Italy. Obviously, these areas completely lie above the eye region.

Minimizing the coefficient of the second fundamental form g along the v direction inside the region of interest, the real v coordinates have been found. In Figure 2 the block diagram of the extraction process is presented.

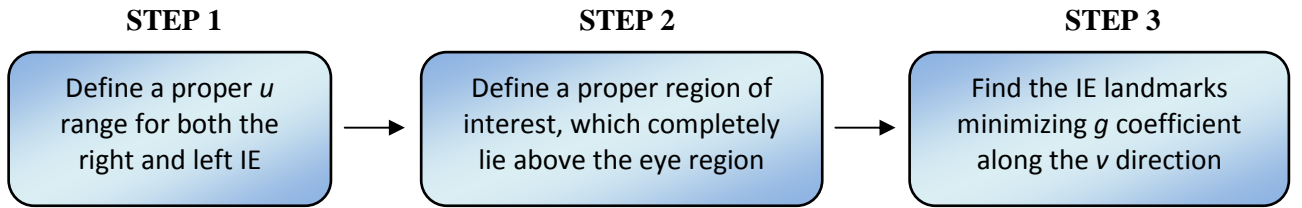


Figure 2. Inner eyebrows identification block diagram.

2.2 OE- outer eyebrow

The *outer eyebrows* are that points which lie at the end of the eyebrows themselves, close to the temples. As happened for the *inner eyebrows* localization, even the extraction of the *outer eyebrows* is done using the coefficient of the second fundamental form g , starting this time from the coordinates of the *exocanthions* previously found.

The process here developed is structured in the same way as the one involving the *inner eyebrows*: firstly, the *outer eyebrows* candidate points are the ones having parametric coordinates, respectively for the right and the left eyebrow, in the ranges:

$$\begin{aligned} u_{OEr} &\in [u_{EXr} - 2, u_{ENr} + 2], & v_{OEr} &\in [v_{EXr} + 5, v_{EXr} + 20], \\ u_{OEl} &\in [u_{EXl} - 2, u_{ENl} + 2], & v_{OEl} &\in [v_{EXl} + 5, v_{EXl} + 20], \end{aligned}$$

where u_{EX} and v_{EX} are the u and v coordinate of the left or right *exocanthions*. Obviously, even these areas completely lie above the eye region.

Minimizing the coefficient g of the second fundamental form along the v direction inside the region of interest, the real v coordinates have been found.

2.3 CH-Cheilion

The *cheilions* are the two points on the outer corners of the mouth, where the outer ends of the upper and lower lips meet. The extraction of these two points is not so simple, due to different conditions that have to be applied in order to properly reduce the region of interest. These are their geometrical features:

1. they belong to the points whose shape index lies in the range corresponding to the surfaces of saddle rut and saddle (Vezzetti *et al.*, 2013);
2. they belong to the points whose curvedness index is positive and greater than a given threshold;
3. the coefficient of the second fundamental form g , going from the left to the right, increases its value for the right *cheilion* and decreases for the left one;
4. the first derivative with respect to the v coordinate is positive for the upper lip while is negative for the lower one. The zero is represented by the line at the end of which the two *cheilions* lie;
5. in a standard pose they lie in the two little hollows close to the mouth, so they are local minimums.

In order to extract them, firstly the algorithm identify a proper region of interest starting from the coordinate of the *subnasal* and of the two *alae*, using the following formulas that come out from geometrical considerations about human faces:

✓ for the right *cheilion*:

$$u_{CHr} \in [u_{ALAr} - 15, u_{ALAr} + 10], \quad v_{CHr} \in [20, v_{SN} - 10],$$

✓ for the left *cheilion*:

$$u_{CHl} \in [u_{ALAl} - 10, u_{ALAl} + 15], \quad v_{CHl} \in [20, v_{SN} - 10],$$

where u_{ALAr} and u_{ALAl} are the u coordinates of the two *alae*, while v_{SN} is the v coordinate of the *subnasal*.

After a preliminary brief geometrical elaboration of the previous region of interest, the algorithm starts working on the curvedness index C , analyzing, for every point belonging to this area, this morphological operator for all the points surrounding it. Then, choosing the values of maximums and minimums, it computes the biggest variation of C within this 3-by-3 sub matrix and compares this value with a proper threshold C_{min} ; if it is lower, the point is erased from the region of interest.

Once this phase is completed, the code starts analyzing the shape index S of all the points which "survived" to the first elimination; again, a proper threshold S_{min} is given. If the analyzed point has a S value less than the threshold, is erased from the region of interest. Then, the algorithm looks for some eventually isolated points, which are not surrounded by any others, and erase them: this operation is preparatory for the last one, in order to give best final results.

Finally, moving from the right to the left for the right *cheilion* and opposite for the left one, the code analyzes all the columns of the region of interest, with the goal to obtain the u coordinates of the two landmarks, identifying the outermost columns which has, at least, one point "survived" to all the elimination processes. The reason why we chose to work this way is given by a preliminary experimental analysis on a sample set of faces.

Lastly, the algorithm analyzes all the points on the left inside the region of interest which have the same u coordinate identified before, in order to obtain the v coordinates of the two *cheilion*. If there is only a single point in the column under analysis, the research is very simple, because this point will be the landmark itself; if otherwise there are two or more points left, the v coordinate of the landmark is obtained as the mean, rounded down, of all the v coordinates of such points. This landmark extraction approach comes out after a deep preliminary experimental work, which showed the method goodness. In Figure 3 the block diagram of the extraction process is presented.

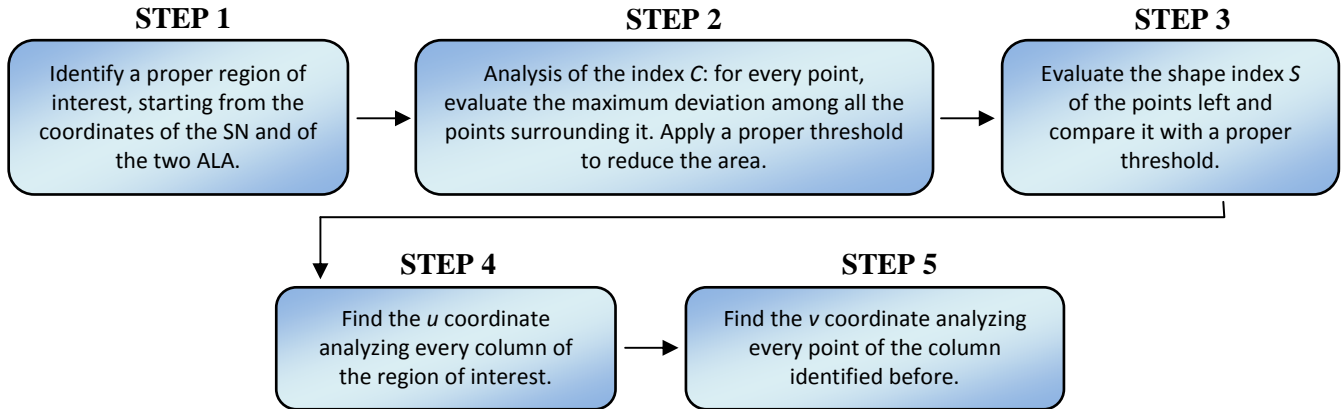


Figure 3. Cheilions identification block diagram.

2.4 LS-Labrum superior

The *labrum superior* is the point which lies on the upper lip of the mouth, approximately in the middle between the two *cheilions*, on that little bump under the subnasal. Its detection is trivial, by using once again the curvedness index C.

Firstly, the u coordinate is manually detected evaluating the mean rounded down of the two u coordinates of the *cheilions*:

$$u_{LS} = \frac{u_{CHl} + u_{CHr}}{2},$$

then, a proper v coordinate range, in which the extraction process is done, is defined:

$$v_{LS} \in [\min(v_{CHr}, v_{CHl}) + 3, v_{SN} - 4].$$

Landmark extraction is achieved with a maximization process algorithm, using a proper threshold C_{min} of the curvedness index C, defined during the preliminary experimental work. In Figure 4 the block diagram of the extraction process is presented.

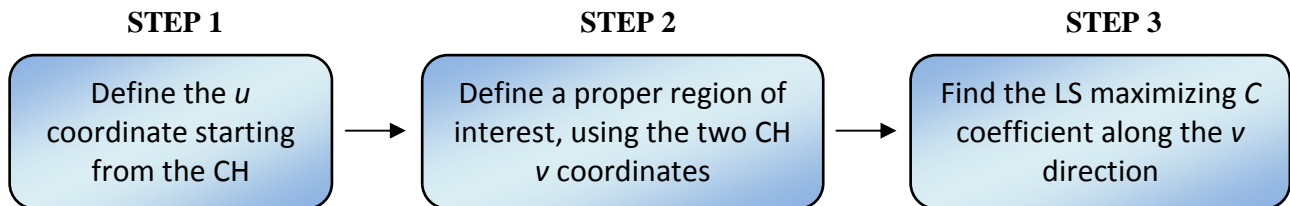


Figure 4. Labrum superior identification block diagram.

2.5 LI-Labrum inferior

The *labrum inferior* is the point which lies on the lower lip of the mouth, approximately in the middle between the two *cheilions*, closer the hollow located above the chin. It is detected, as happened for the *labrum superior*, by maximizing the curvedness index C on a proper region of interest, defined this way:

$$u_{LI} = \frac{u_{CHl} + u_{CHr}}{2} = u_{LS} ,$$

$$v_{LI} \in [\min(v_{CHr}, v_{CHl}) - 7, v_{LS} - 9] .$$

The landmark extraction is made using a proper threshold C_{min} of the curvedness index C , defined during the preliminary experimental work.

2.6 Other eyebrows points

Although they are not properly landmarks, some other eyebrow points have been detected, in order to best identify their real shape. To do this, five inner points among the inner eyebrows and the outer eyebrows have been extracted. As happened for the IE and the OE, even all these new eyebrows points have been obtained through the minimization of the coefficient of the second fundamental form g along the v coordinate, keeping constant the u parameter. Firstly, it has been evaluated the u distance, named Δu , between two generic consecutive point:

✓ for the right eyebrow:

$$\Delta u_r = \frac{(u_{ENr} + 5) - u_{EXr}}{n - 2} ,$$

✓ for the left eyebrow:

$$\Delta u_l = \frac{u_{EXl} - (u_{ENl} + 5)}{n - 2} ,$$

where u_{EN} and u_{EX} are the u coordinates of, respectively, the left and right *endocanthions* and *exocanthions*. After this, the u coordinate of all the five points have been evaluated, using the following formula:

$$u_{Ei} = u_{ENr/l} \pm i\Delta u_{r/l} \quad 1 \leq i \leq 5 ,$$

where the "+" sign is used for the left eyebrow, while the "-" sign is used for the right one. After this operation, the initial v coordinate range of all the five points has been manually set, exactly as it has been done for both the inner and outer eyebrows:

$$v_{EN} + 5 \leq v_{E1, \dots, 5} \leq v_{EN} + 20 .$$

Lastly, minimizing the coefficient of the second fundamental form g along the v direction, the real v coordinates have been found.

It has been observed that, in some of the shells belonging to our database, slight identification problems rose, obtaining bad or, in the worst case scenario, wrong results. In order to have better quality results, some additional strings have been introduced at the end of the algorithm which could erase some previous points overwriting them with new ones, obtained working only on geometrical features of the eyebrows.

For example, the algorithm checks out if there are big differences between the v coordinates of all the eyebrows points: it could be well known that the eyebrows themselves have an ideal arch form, split across the whole eye region, so high v coordinates variations could not be noticed. Introducing proper thresholds, if the algorithm finds out that one or more points have, compared to all the others, too high relative differences Δv among their v coordinates, it erases the incorrect points and replaces them with new ones which satisfy threshold conditions.

2.7 Other mouth points

Similarly to the other eyebrow points, six points on the upper lip and six other points on the lower one have been extracted. All the twelve points are identified in the same way, i.e. by maximizing, as happened for the LS and LI landmarks, the curvedness index C using a proper threshold C_{min} .

Named $M1$, $M2$, and $M3$ the points that can be found starting from the *labrum superior* and going to the right *cheilion*; $M4$, $M5$, and $M6$ the ones going from the *labrum superior* to the left *cheilion*; $M7$, $M8$, and $M9$ the ones from the *labrum inferior* and going to the right *cheilion*; $M10$, $M11$, and $M12$ the ones going from the *labrum inferior* to the left *cheilion*, we can define their u coordinates:

$$\begin{cases} u_{M1} = u_{LS} + \Delta u = u_{M7} , \\ u_{M2} = u_{M1} + \Delta u = u_{M8} , \\ u_{M3} = u_{M2} + \Delta u = u_{M9} , \\ u_{M4} = u_{LS} - \Delta u = u_{M10} , \\ u_{M5} = u_{M4} - \Delta u = u_{M11} , \\ u_{M6} = u_{M5} - \Delta u = u_{M12} , \end{cases}$$

where Δu is computed this way:

$$\Delta u = \frac{u_{CHl} - u_{CHr}}{8} .$$

The v coordinate range, in which the maximization process is done, is computed this way:

- ✓ for the points belonging to the upper lip:

$$\min(v_{CHr} , v_{CHl}) + 3 \leq v_{M1,...,6} \leq v_{SN} - 4 ;$$

- ✓ for the points belonging to the lower lip:

$$\min(v_{CHr} , v_{CHl}) - 7 \leq v_{M7,...,12} \leq v_{LS} - 9 .$$

It has been observed that, in some of the shells belonging to our database, slight identification problems rose, obtaining bad or, in the worst case scenario, wrong results. In order to have better quality results, some additional strings have been introduced at the end of the algorithm which could erase some previous points overwriting them with new ones, obtained working only on geometrical features of the mouth and introducing proper thresholds.

3. Experimental validation

The meshes used in the experiment are obtained with two different techniques. For some of them we used a Minolta Vivid 910 3D laser scanner, for the others a photogrammetric system composed by four Canon EOS 1100D cameras. The database is composed by nine Caucasian

subjects, both male and female, from 20 up to 40 years old, each one portrayed in seven different facial expressions: serious (standard expression), anger, disgust, fear, happiness, sadness, and surprise, reaching a total of 63 different meshes. This choice is based on the theory of “basic emotions” of Ekman (Ekman, 1970; Ekman and Keltner, 1997), whose studies showed that these were the six main emotional expressions. Due to some problems inherent to shells themselves, the database has been reduced to 58 meshes.

Furthermore, since we have considered that 58 shells obtained only through our physical tools could not be sufficient for testing our developed algorithm, we have tested it also on public Bosphorus database. This database is a variegated and diversified set of faces coming from 105 people, including partial shells, with holes, and non-frontal faces. Among them, we chose 54 frontal-view faces coming from nine people performing six facial expressions each (Ekman's ones, excluding straight one), in accordance with our database. So, the algorithm has been tested on 112 shells overall.

Starting from the previous Matlab® algorithm of eyes and nose landmarks extraction (Vezzetti *et al.*, 2013), a new part has been added in order to automatically compute the missing information about eyebrows and mouth. In the following paragraphs the results of the landmark extraction procedures will be presented using one face as sample. The graphical behaviour of the descriptors, with the corresponding landmark neighbourhood reduction, is shown on the left side, while the landmarks detected are shown on the right one.

3.1 IE-*Inner eyebrow*

The interesting areas of the two *inner eyebrows* are firstly reduced by selecting the points which lie in a proper region of interest obtained starting from the coordinate of the two *endocanthions*. Then, among these points, the *inner eyebrows* are selected through a minimization of the coefficient of the second fundamental form g along the v direction (Fig. 5).

The theoretical localization process seems to be reflected in the application. The algorithm elaborated for the extraction gives the points expected. Similarly to this landmark, the extraction of the other points showed that the theoretical processes are reflected in the application.

3.2 OE-*Outer eyebrow*

The process of the extraction of the *outer eyebrows* consists, as happened for the inner eyebrow extraction, in minimizing the coefficient of the second fundamental form g inside a proper region of interest, evaluated this time respect to the *exocanthions* (Fig. 6).

Although they are less easily identifiable than the *IE*, due to their external positions on the face, results seem to be satisfactory for all the faces inside the database.

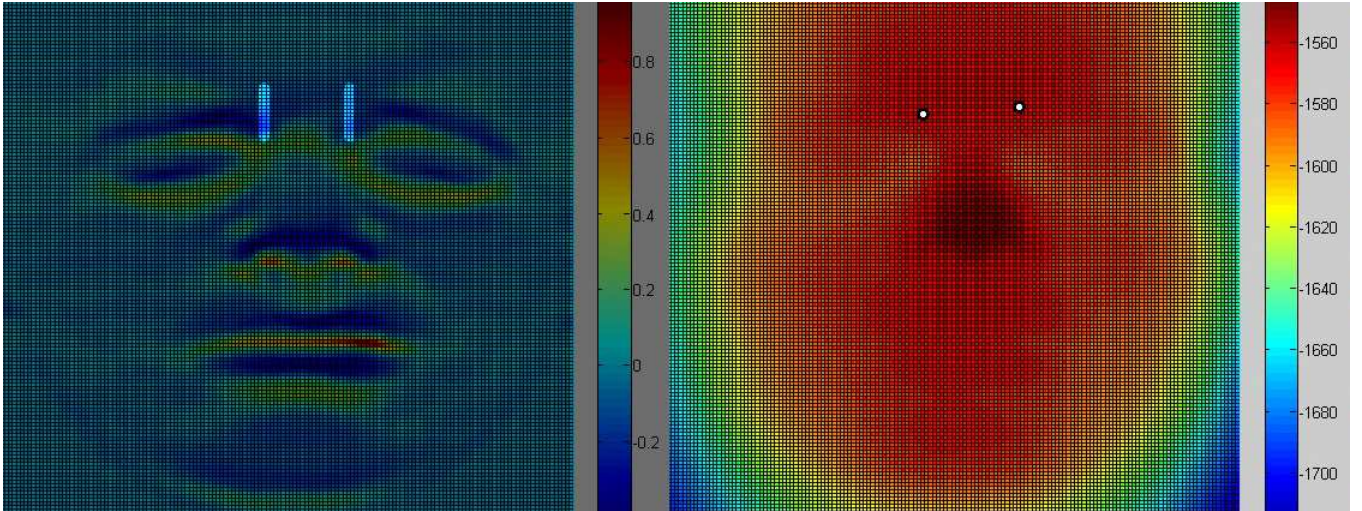


Figure 5. (left) the coefficient of the second fundamental form g : the brightest areas are the ones in which the minimization is done; (right) the *inner eyebrows* (in white) extracted.

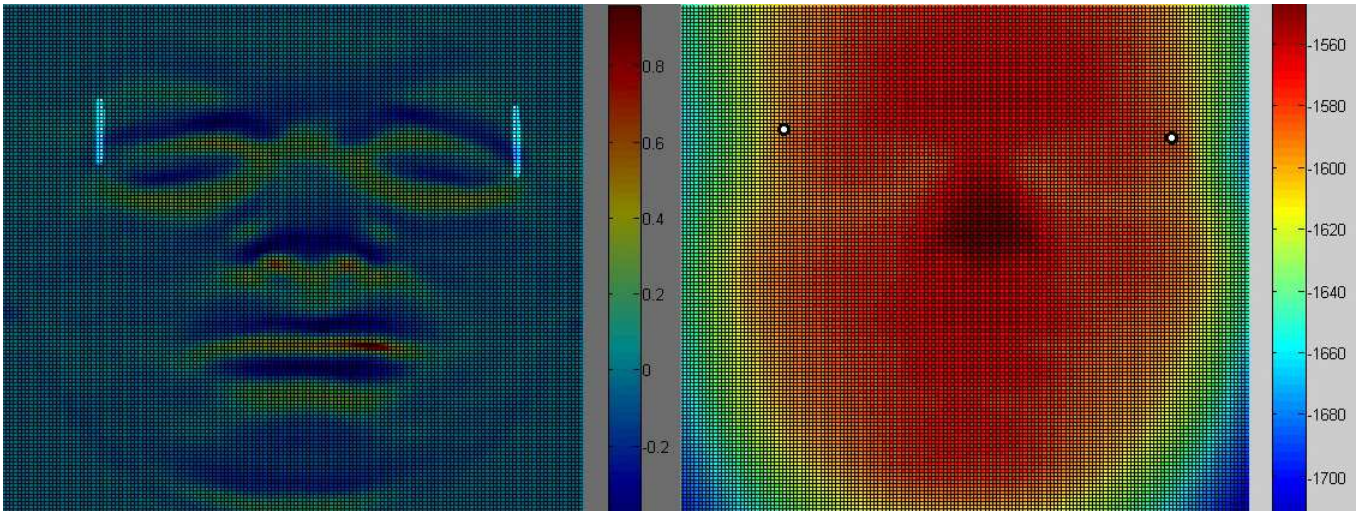


Figure 6. (left) the coefficient of the second fundamental form g : the brightest areas are the ones in which the minimization is done; (right) the *outer eyebrows* (in white) extracted.

3.3 CH-Cheilion

Unlike the other two eyebrow landmarks, the *cheilion* has been harder to identify, due to the various conditions applied in order to properly reduce the region of interest. The algorithm works with different morphological indexes: once the region of interest is defined, two conditions, on C and on S indexes, are firstly applied, then the *cheilions* are extracted among the points which "survived" to the elimination process. In Figures 7-9 the steps done by the algorithm to detect these two landmarks are shown.

Results on the whole set of faces show the method goodness, obtaining most of times correct landmarks or, in the worst cases, identifying at least the real region in which the landmarks lie.

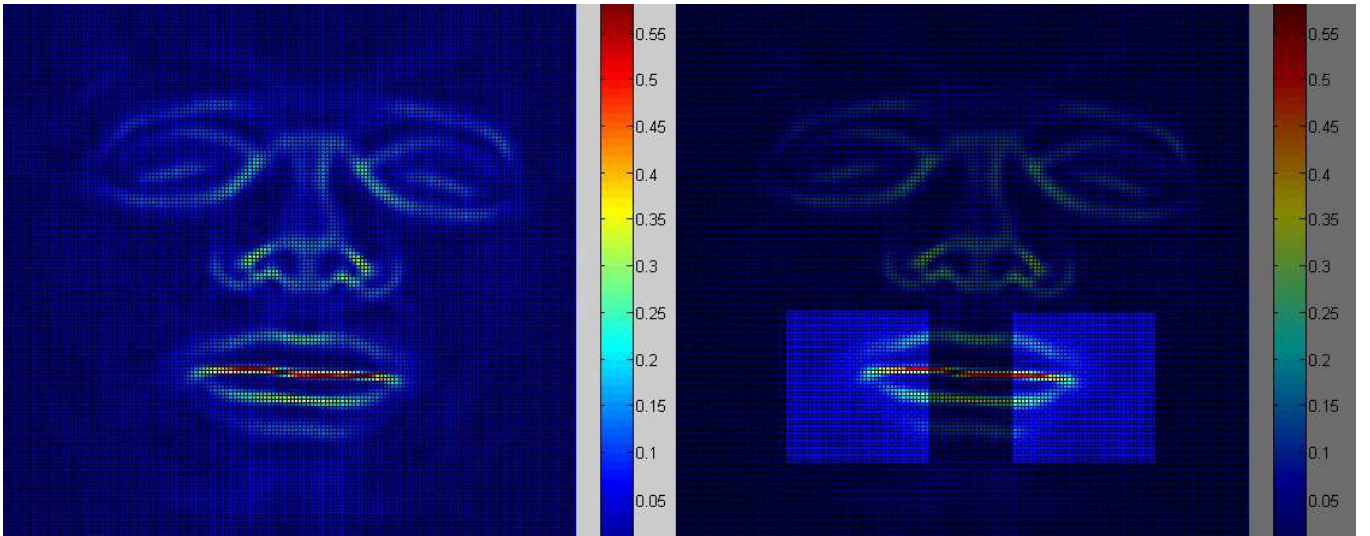


Figure 7. (left) the curvedness index C ; (right) the initial the region of interest (the bright area).

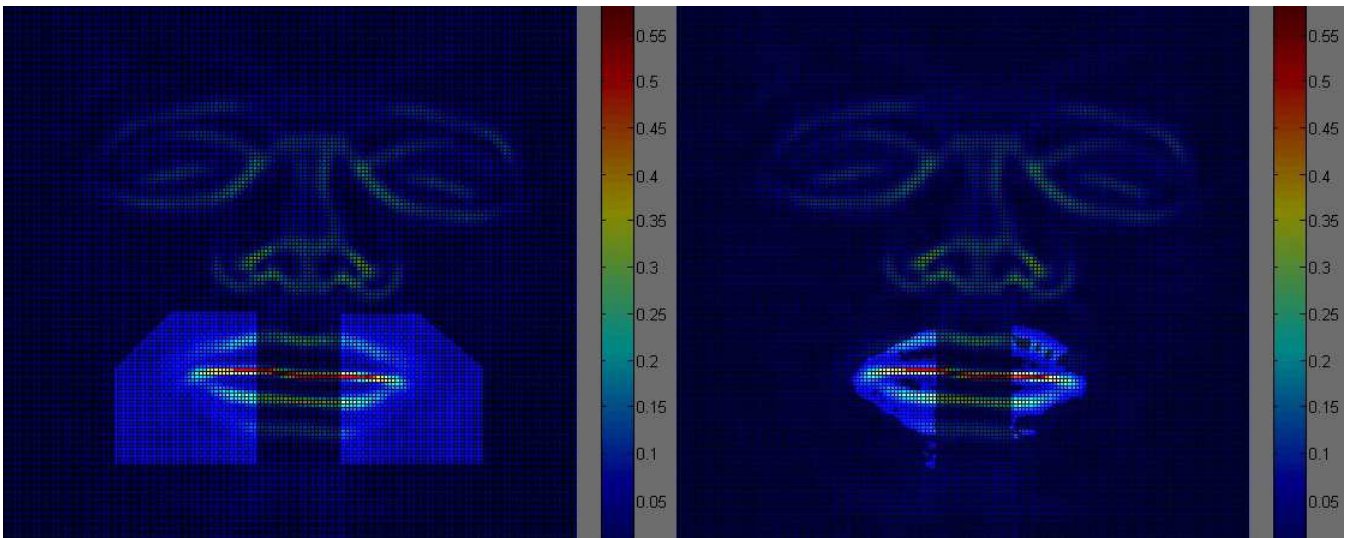


Figure 8. (left) the region of interest after the first geometrical elaboration; (right) the reduced area after the application of the C condition.

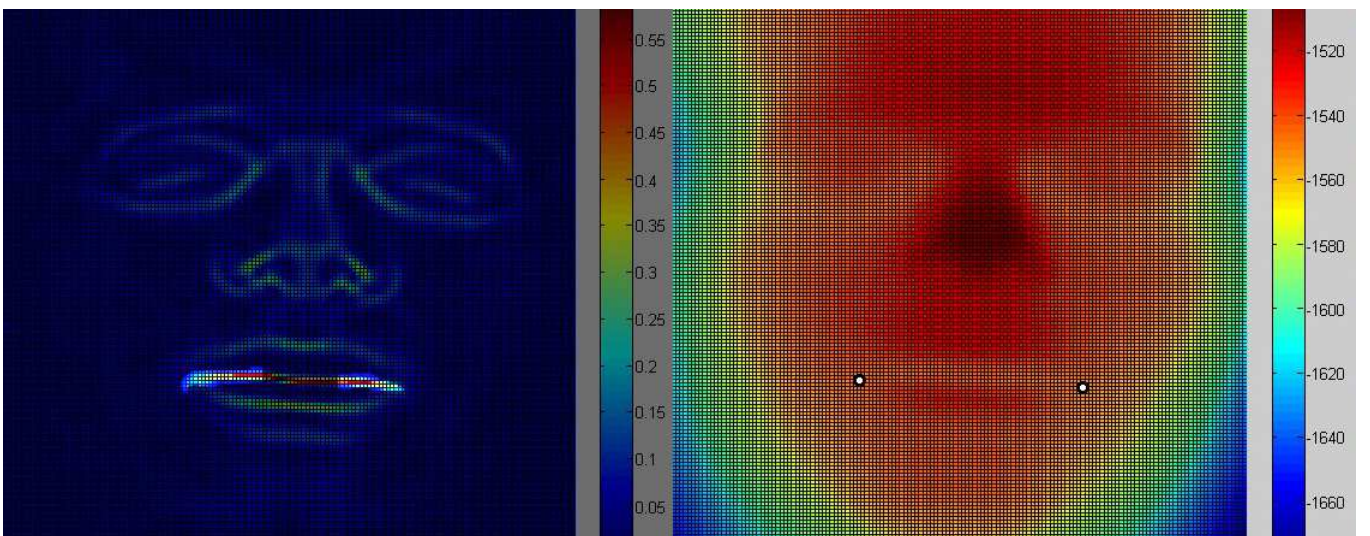


Figure 9. (left) the region of interest after the application of the S condition; (right) the two *cheilions* extracted (in white).

3.4 LS-Labrum superior

The area of interest of the *labrum superior* is firstly defined by selecting the points which lie in a proper region of interest starting from the coordinates of the two *cheilions* and of the *subnasal*. Then, among these points, the *labrum superior* is extracted through a maximization of the curvedness index C along the v direction. In Figure 10 the steps done by the algorithm to detect the *labrum superior* are shown.

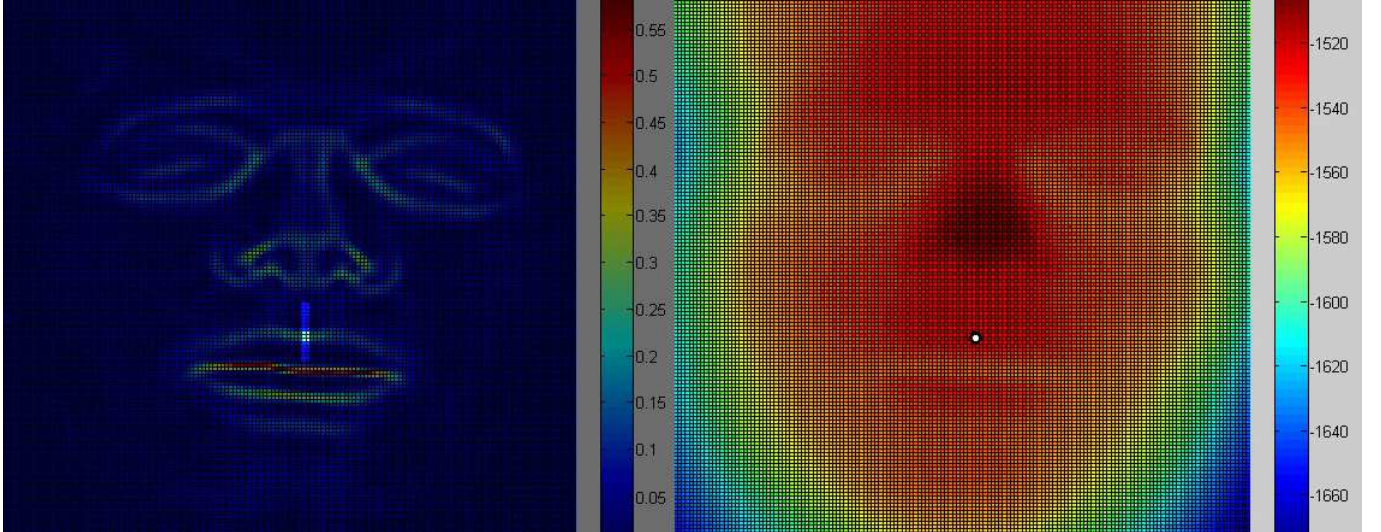


Figure 10. (left) maximization of the curvedness index C inside the region of interest (the brightest area); (right) the *labrum superior* extracted (in white).

3.5 LI-Labrum inferior

Similarly to the LS landmark, the *labrum inferior* is obtained by maximizing the curvedness index C among the points which lie in a proper region of interest, defined starting from the coordinates of the two *cheilions* and of the *labrum superior*. In Figure 11 the steps done by the algorithm to detect the *labrum inferior* are shown.

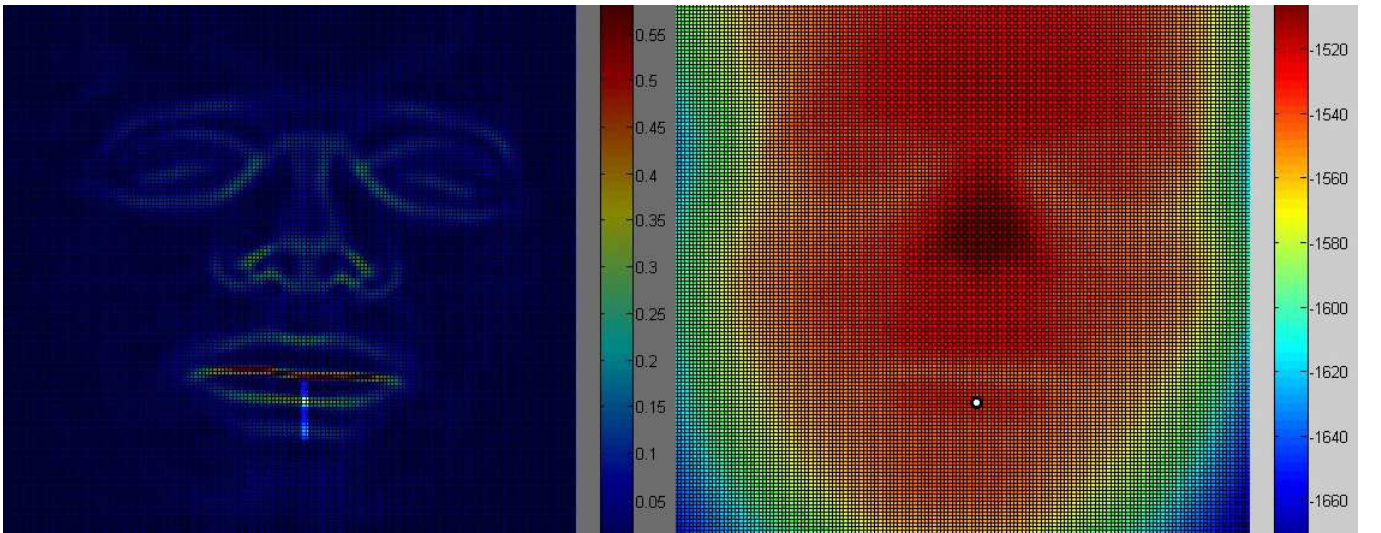


Figure 11. (left) maximization of the curvedness index C inside the region of interest (the brightest area); (right) the *labrum inferior* extracted (in white).

3.6 Other eyebrows points

As happened for the IE and the OE, also these eyebrows points have been obtained through the minimization of the coefficient of the second fundamental form g along the v direction, keeping constant the u parameter. We chose to extract this point in order to give a better shape to the eyebrow itself, identifying five inner points, as shown in Figure 12.

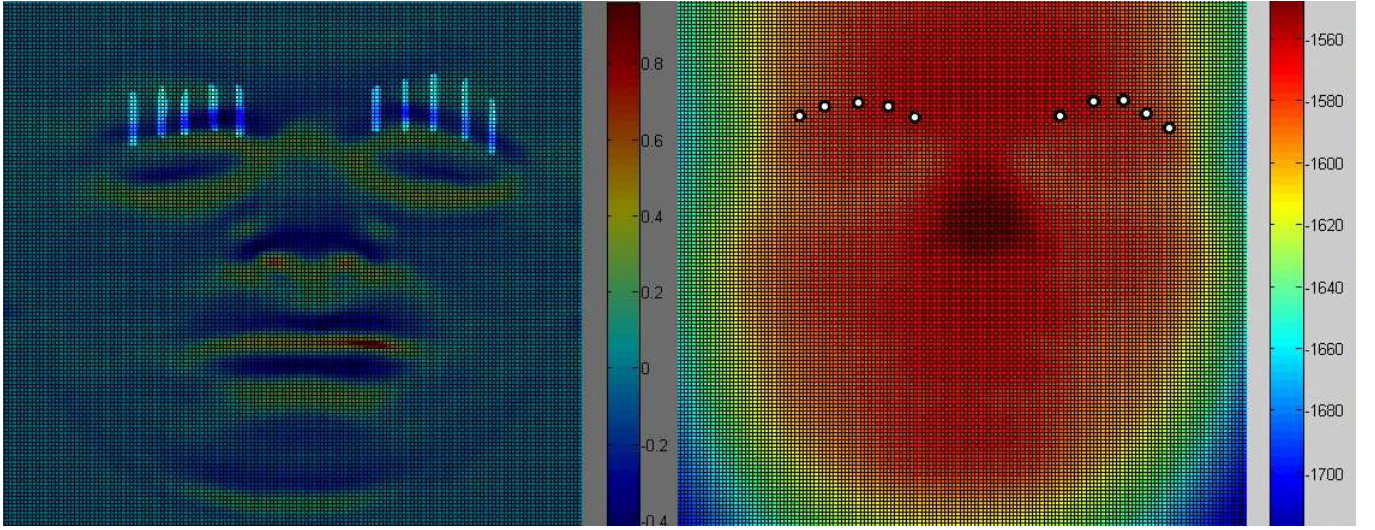


Figure 12. (left) minimization of the coefficient of the second fundamental form g inside the region of interest (the brightest areas); (right) the six extracted points (in white).

3.7 Other mouth points

Applying the same considerations made for the eyebrows, for the mouth we decided to extract some inner points with the goal to best identify the shape of the mouth itself. Six points are obtained on the upper lip and six others on the lower one. In Figure 13 the steps done by the algorithm to detect all these points are shown.

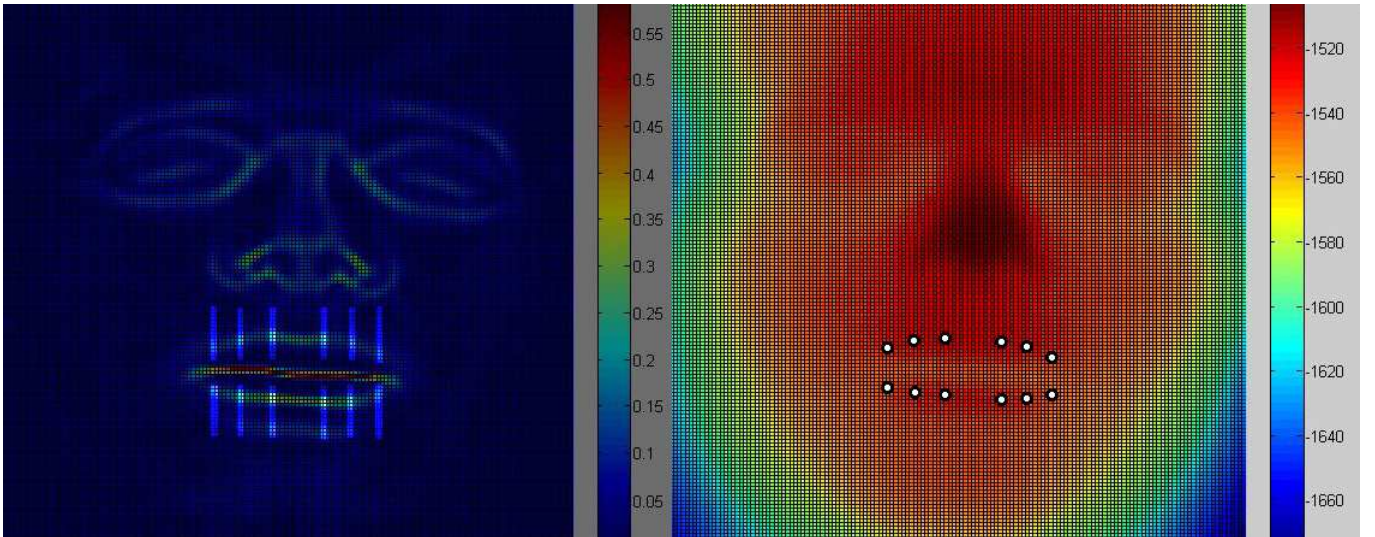


Figure 13. (left) maximization of the curvedness index C inside the region of interest (the brightest areas); (right) the twelve points extracted (in white).

4. Results

The algorithm implemented has a global computing time of about two seconds. Figure 14 shows the results of the algorithm performed on the same face portrayed in all the seven different

facial expressions. Figure 15 presents nine different subjects faces performing the same facial expression, the fear one.

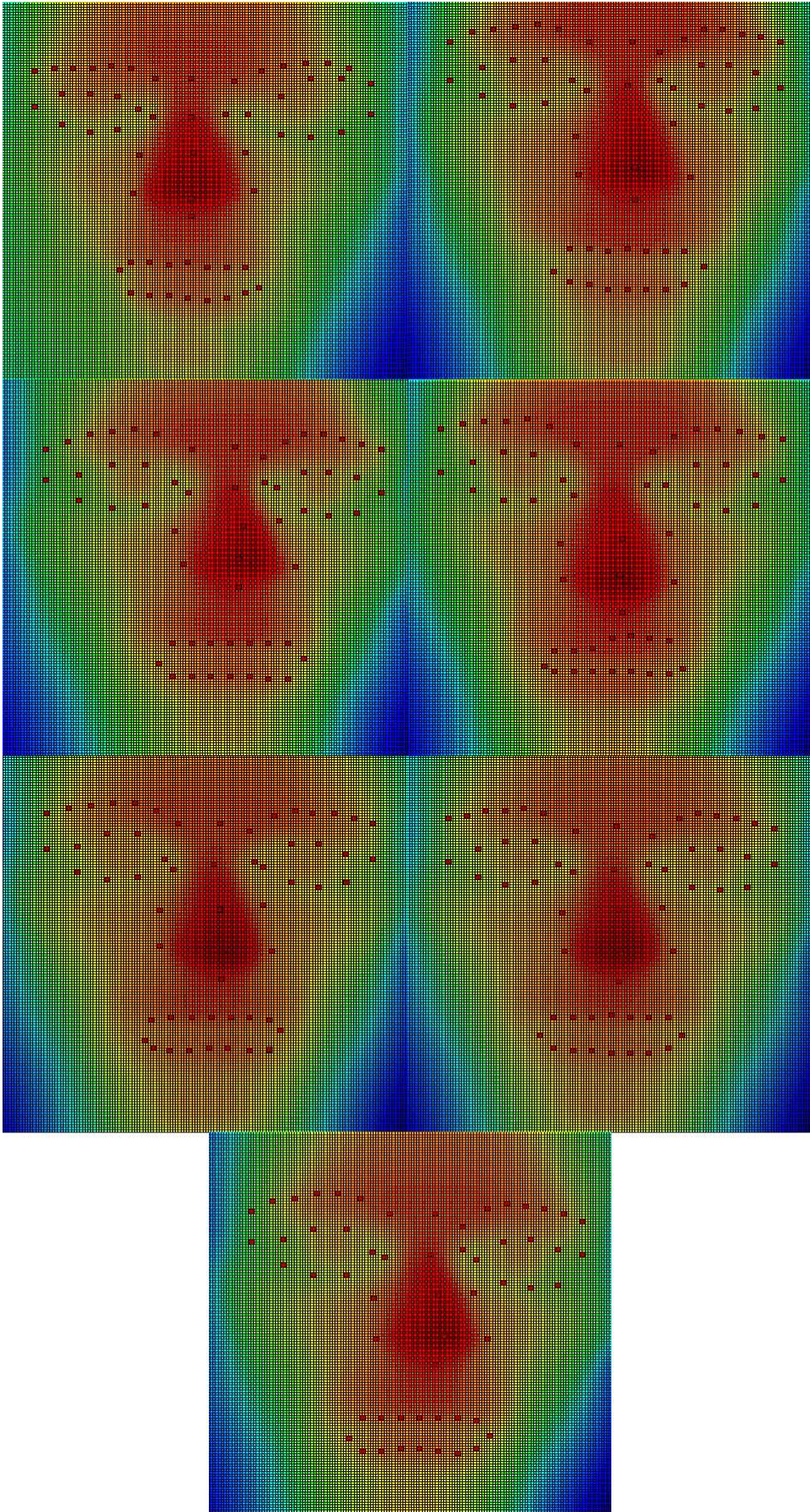


Figure 14. From left to right, from the top to the bottom: disgust, happiness, fear, anger, serious, surprise, and sadness expressions.

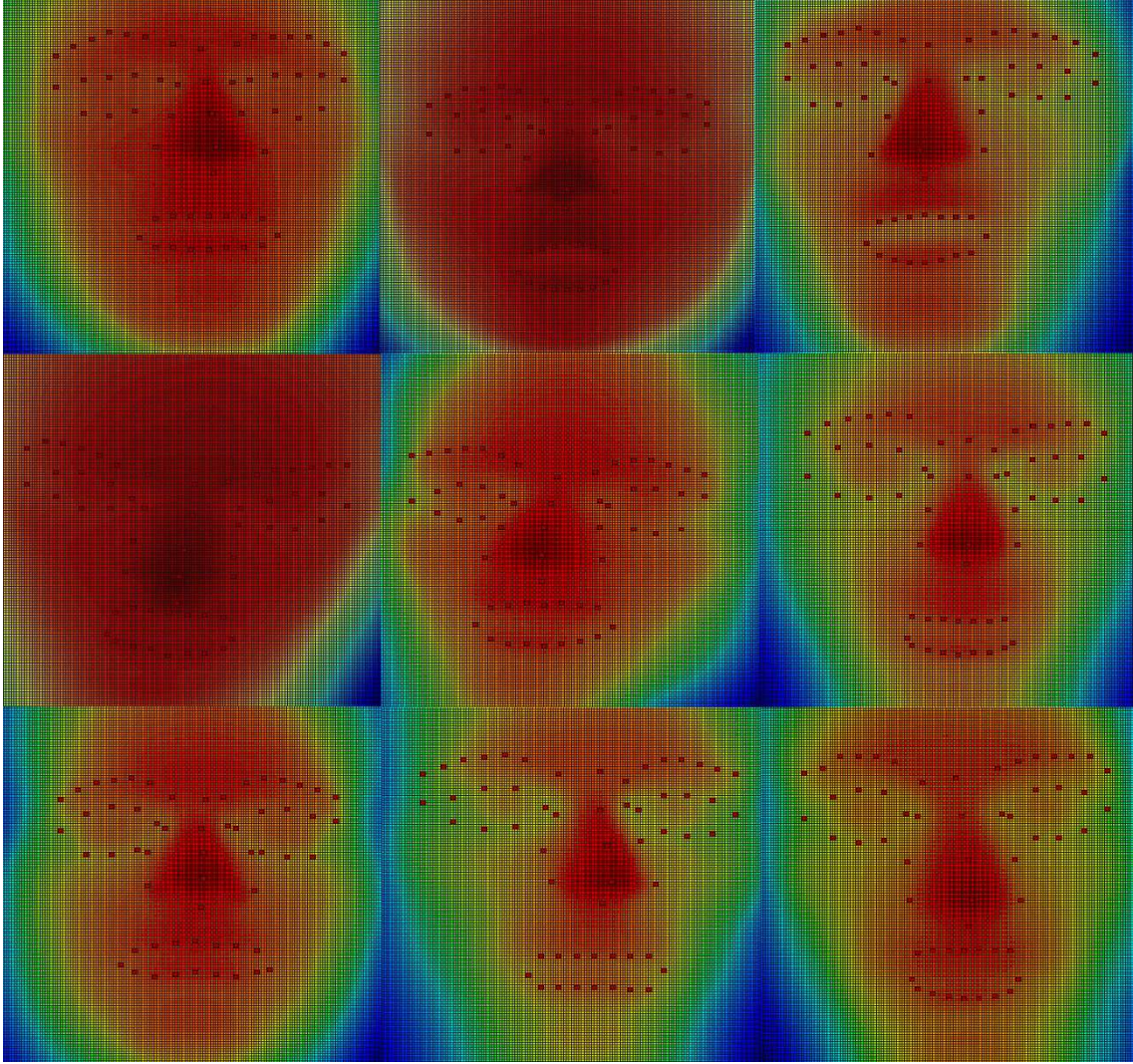


Figure 15. Nine subjects performing fear expression.

All the landmarks detected are here represented by brown rectangles inside the figures, which are introduced only for visibility reasons; the real landmarks lie at the top-left bottom pixel of those rectangles. As can be noticed, all the landmarks of all the faces are properly detected: most of them are identified in an accurate position, however some of them have little positioning errors, which are now computed in a more rigorous way.

In order to verify the position accuracy of all the extracted landmarks, a statistical study has been performed. Firstly, landmarks have been hand-detected from a plastic surgeon on these 112 faces, so that a comparison between these landmarks and the extracted ones could be realized. Only the principal landmarks have been considered, omitting all the other points extracted only to best define the nose, the eyes, the eyebrows, and the mouth shapes.

Subsequently, Euclidean distances between the real landmarks and the respective points given by the algorithm have been computed: to do this, it has been previously necessary to evaluate the real dimension in the international metric system of one grid unit, remembering that our grid has 150 units for both the u and the v directions.

To do this, the idea of a face bounding box can be very useful: a bounding box is an invisible 3D rectangular box in which the face is completely inscribed, as shown in Figure 16.

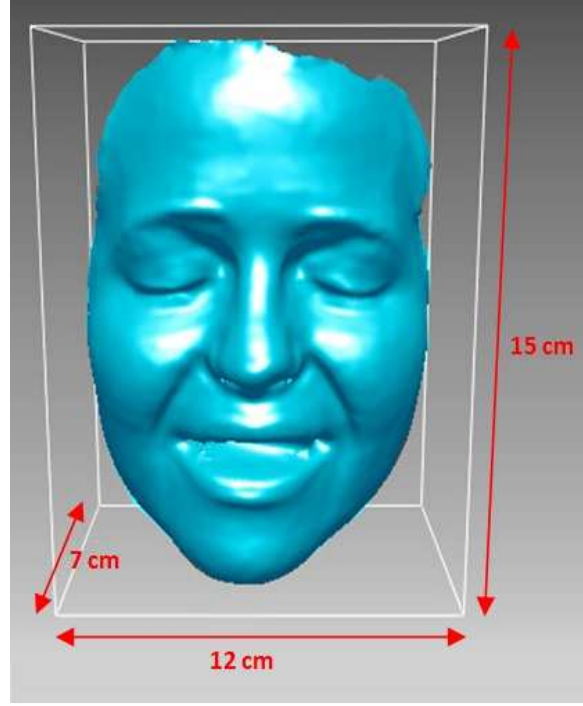


Figure 16. The face bounding box and its sizes.

The sides of this box have a standard length: it can be noticed that the face width is about 12 cm, the face height is about 15 cm and the face depth is about 7 cm. Using these face bounding box sizes and knowing that our mesh grids are square matrices of 150 X 150 elements, applying the following proportions it can be easy to compute the length of one unit grid:

✓ for the u direction:

$$150 \text{ units} : 120 \text{ mm} = 1 \text{ units} : x \text{ mm} \rightarrow 1 \text{ unit} = 0.8 \text{ mm};$$

✓ for the v direction:

$$150 \text{ units} : 150 \text{ mm} = 1 \text{ units} : x \text{ mm} \rightarrow 1 \text{ unit} = 1 \text{ mm}.$$

It is now possible to evaluate the real distances, in mm, between the real landmarks and the ones extracted by the algorithm, applying the Pythagorean Theorem this way:

$$d = \sqrt{[0.8 (u_r - u_e)]^2 + [1 (v_r - v_e)]^2},$$

where u_r , u_e , v_r , and v_e are, respectively, the u and v coordinates of the real and the extracted landmarks. Two different analyses have been performed: the first in order to evaluate the scanning goodness of the single person set of faces that he/she performed, the second in order to evaluate the algorithm accuracy in extracting the different landmarks analyzing; for each one, the position results among the whole database.

According to the maxillofacial plastic surgeon, a proper threshold value for evaluating the distance error between real and extracted landmarks could be 4 mm, thus it can be said that all the identified points which have a deviation less or equal to this value are accurately-identified points.

Sample mean μ and the sample variance σ are calculated among the whole set of their position errors. In Figure 17 the graphs of the sample mean and of the standard deviation, computed as $\pm\sqrt{\sigma}$, are reported.

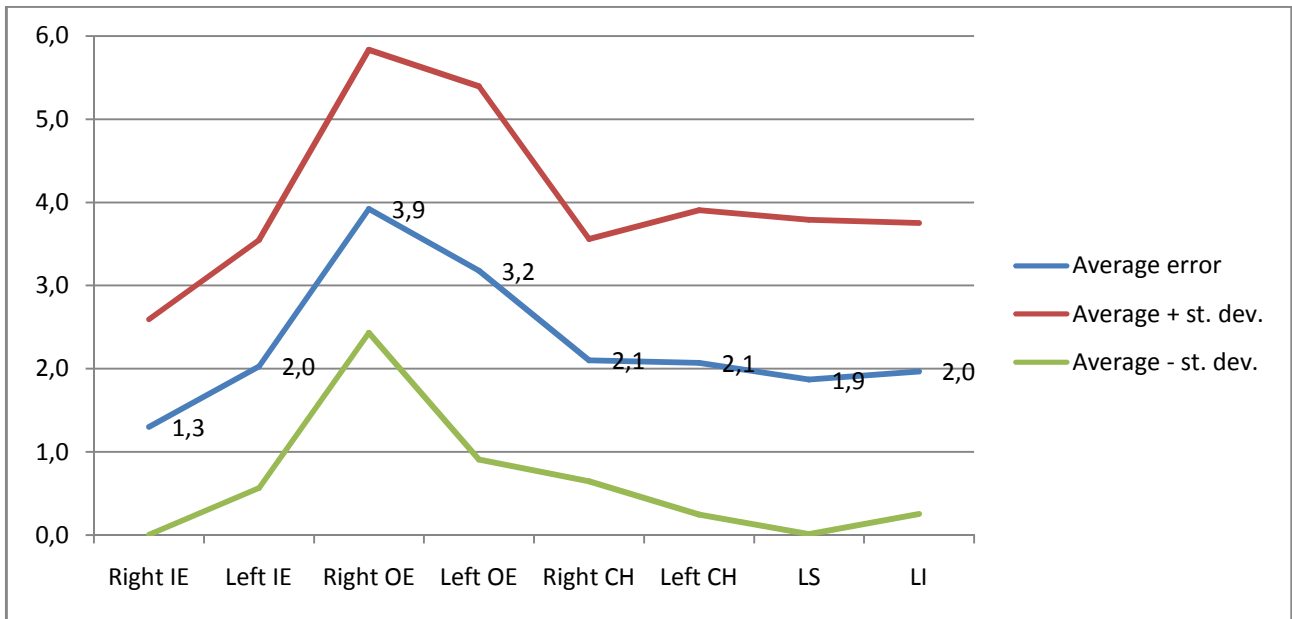


Figure 17. Mean errors and standard deviations for each landmark. The y-axis and the numerical values of the blue average error line are in mm.

Means belong to the range [1.3, 3.9]. The worst result concerns the right *outer eyebrow* points, while the best one for the right *inner eyebrow*. Means are less than 4 mm, so the threshold set by the surgeon has been respected. To better see the extraction accuracy of each point, a frequency analysis above the whole set of 896 landmarks (= 8 landmarks for 112 shells) has been done, as shown in Figure 18.

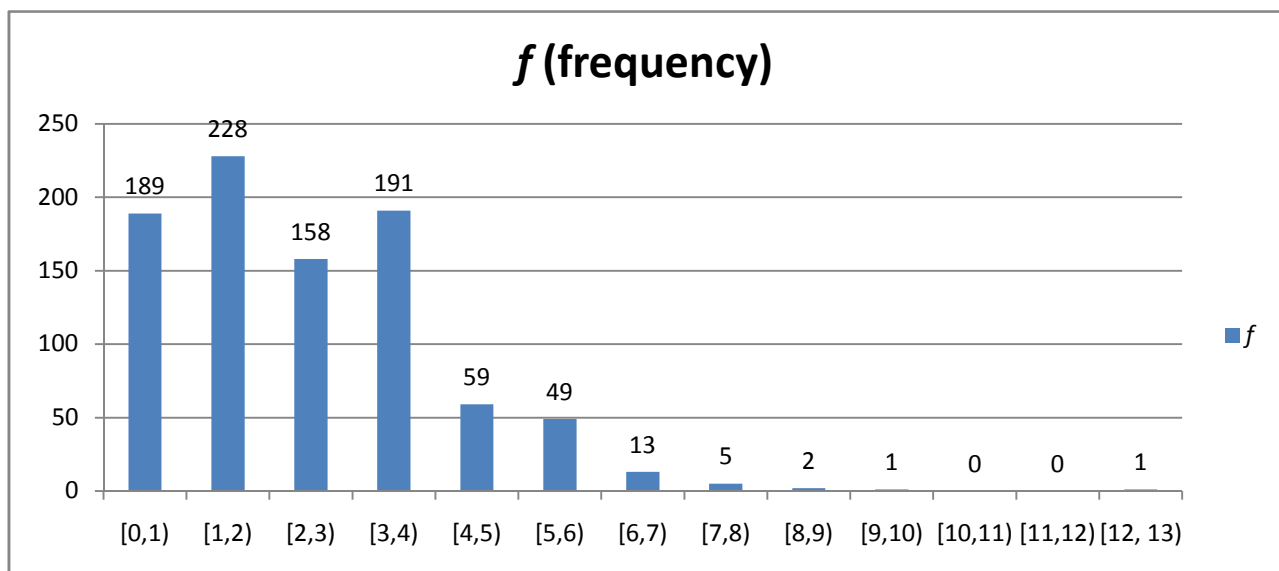


Figure 18. Distribution of all the extracted landmarks. The height of each column is the number of landmarks with an extraction error in the range at the x-axis of its column. For instance, there are 189 landmarks having an error in the range [0,1). The unit is mm.

The graphs show how most of the points are extracted with an error less than the threshold value. Using the data provided in the previous sections, it is possible to confirm the global goodness of the method. In particular, 766 out of 896 landmarks have a position error less than the 4 mm threshold, so about 85% of them have been correctly extracted.

5. Conclusion

Face is the locus of expressions, emotions, and social relationships. Its technical study is important for supporting specialists, surgeons, and physicians in their work. That is why it is important to quantify face shape analyzing it in various positions and expressions. To this end, different three-dimensional approaches could be implemented. In this work, a differential-geometry-based method is proposed as a continuation of our previous 3D landmarking study, with the purpose of extracting some missing landmarks belonging to mouth and eyebrows: *chelions*, *labrum superior*, *labrum inferior*, *inner eyebrows*, and *outer eyebrows*.

A geometric description and outline of the eight new soft-tissue landmarks is provided, then a Matlab® algorithm is implemented to extract them. This algorithm has been tested both on our database of faces and on the public Bosphorus one. The results obtained were shown to an expert, a maxillofacial surgeon, who confirmed their correctness, giving a proof of the goodness of our extraction code and accuracy of results.

References

- Alker, M., Frantz, S., Rohr, K., and Stiehl, H.S., 2001. Improving the Robustness in Extracting 3D Point Landmarks from 3D Medical Images Using Parametric Deformable Models. *Medical Image Computing and Computer-Assisted Intervention* 2208, 582-590.
- Alyüz, N., Gökberk, B., Dibeklioglu, H., Savran, A., Salah, A.A., Akarun, L., and Sankur, B., 2008. 3D Face Recognition Benchmarks on the Bosphorus Database with Focus on Facial Expressions. *Biometrics and Identity Management, Lecture Notes in Computer Science* 5372, 57-66.
- Calignano, F., Vezzetti E., 2010. Soft tissue diagnosis in maxillofacial surgery: A preliminary study on three-dimensional face geometrical features-based analysis. *Aesthetic Plastic Surgery* 34 (2), 200-211.
- Creusot, C., Pears, N., and Austin, J., 2013. A Machine-Learning Approach to Keypoint Detection and Landmarking on 3D Meshes. *International Journal of Computer Vision* 102(1-3), 146-179.
- D'Hose, J., Colineau, J., Bichon, C. and Dorizzi, B., 2007. Precise Localization of Landmarks on 3D Faces using Gabor Wavelets. *First IEEE International Conference on Biometrics: Theory, Applications, and Systems*, 1-6.
- Dibeklioglu, H., Salah, A.A., and Akarun, L., 2008. 3D Facial Landmarking under Expression, Pose, and Occlusion Variations. *2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, 1-6.
- Ekman, P., 1970. Universal Facial Expressions of Emotions. *California Mental Health Research Digest* 8(4), 151-158.
- Ekman, P. and Keltner, D., 1997. *Facial Expressions of Emotions*, Lawrence Erlbaum Associates Publisher, Mahwah, New Jersey.

- Frantz, S., Rohr, K. and Stiehl, H.S., 1998. Multi-Step Procedures for the Localization of 2D and 3D Point Landmarks and Automatic ROI Size Selection. *Computer Vision* 1406, 687-703.
- Frantz, S., Rohr, K. and Stiehl, H.S., 1999. Improving the Detection Performance in Semi-automatic Landmark Extraction. *Medical Image Computing and Computer-Assisted Intervention* 1679, 253-262.
- Frantz, S., Rohr, K. and Stiehl, H.S., 2000. Localization of 3D Anatomical Point Landmarks in 3D Tomographic Images Using Deformable Models. *Medical Image Computing and Computer-Assisted Intervention* 1935, 492-501.
- Frantz, S., Rohr, K. and Stiehl, H.S., 2005. Development and validation of a multi-step approach to improved detection of 3D point landmarks in tomographic images. *Image and Vision Computing* 23(11), 956-971.
- Perakis, P., Palassis, G., Theoharis, T. and Kakadiaris, L.A., 2009. Automatic 3D Facial Region Retrieval from Multi-pose Facial Datasets. *Eurographics Workshop on 3D Object Retrieval*.
- Romero, M. and Pears, N., 2009a. Landmark Localization in 3D Face Data. 6th IEEE International Conference on Advanced Video and Signal Based Surveillance, 73-78.
- Romero, M. and Pears, N., 2009b. Point-pair descriptors for 3D facial landmark localization. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, 1-6.
- Ruiz, M.C. and Illingworth, J., 2008. Automatic landmarking of faces in 3D - ALF3D. 5th International Conference on Visual Information Engineering - IEEE Conferences, 41-46.
- Salah, A.A. and Akarun, L., 2006a. 3D Facial Feature Localization for Registration. *Multimedia Content Representation, Classification and Security* 4105, 338-345.
- Salah, A.A. and Akarun, L., 2006b. Gabor Factor Analysis for 2D+3D Facial Landmark Localization. *IEEE 14th Signal Processing and Communications Applications*, 1-4.
- Salah, A.A., Çinar, H., Akarun, L., and Sankur, B., 2007. Robust Facial Landmarking for Registration. *Annals of Telecommunications* 62 (1-2), 83-108.
- Sang-Jun, P. and Dong-Won, S., 2008. 3D face recognition based on feature detection using active shape models. *International Conference on Control, Automation and Systems - IEEE Conferences*, 1881- 1886.
- Vezzetti, E. and Marcolin, F., 2012. Geometry-based 3D face morphology analysis: soft-tissue landmark formalization. *Multimedia Tools and Applications*, DOI: 10.1007/s11042-012-1091-3.
- Vezzetti, E., Marcolin, F., and Stola, V., 2013. 3D human face soft tissues landmarking method: An advanced approach. *Computers in Industry*, in press <http://dx.doi.org/10.1016/j.compind.2013.04.006>.
- Wörz, S. and Rohr, K., 2005. Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. *Medical Image Analysis* 10(1), 41-58.
- Zhang, X., Wang, Y., & Pan, G., 2013. 3D facial landmark localization via a local surface descriptor HoSNI. *Intelligent Science and Intelligent Data Engineering*, Springer Berlin Heidelberg, 313-321.